Molecular Breeding **3:** 257–268, 1997.

Mapping and QTL analysis of horticultural traits in a narrow cross in cucumber (*Cucumis sativus* L.) using random-amplified polymorphic DNA markers

Felix C. Serquen[1], J Bacher[2] & JE Staub[3,*]

[1]*Department of Horticulture, 1575 Linden Dr., University of Wisconsin, Madison, WI 53706, USA;* [2]*Promega Corporation, 2800 Woods Hollow Rd., Madison, WI 53711, USA;* [3]*USDA-ARS, Vegetable Crops Unit, Department of Horticulture, 1575 Linden Dr., University of Wisconsin, Madison, WI 53706, USA (*author for correspondence)*

Received 16 September 1996; accepted in revised form 6 March 1997

Key words: determinate, genetic markers, multiple lateral branching, plant architecture, quantitative trait loci, sex expression, cucumber, RAPD

Abstract

An 80-point genetic map [77 random-amplified polymorphic DNAs (RAPD), *F* (female sex expression), *de* (determinate), and *ll* (little leaf)] was constructed from a narrow cross in cucumber using the determinate, gynoecious, standard-sized leaf line G421 and the indeterminate, monoecious, little leaf line H-19. The map defined nine linkage groups and spanned ca. 600 cM with an average distance between markers of 8.4 ± 9.4 cM. The RAPD loci BC-551 and BC-592 were found to flank *ll* at 3.4 and 12.2 cM, respectively. The locus OP-L18-2 was linked (16 cM) to *de*, and the *F* locus was flanked by markers at 44 and 31 cM. One-hundred F_3 families were used to identify quantitative trait loci (QTL) for sex expression, main stem length, number of lateral branches, days to anthesis, fruit number and weight, fruit length and diameter, and fruit length: diameter ratio in two replicated test locations (Wisconsin and Georgia). QTL on linkage group B explained major portions (R^2 = ca. 2 to 74%) of the variation observed for sex expression, main stem length, lateral branch number, and fruit diameter (LOD = 2.1 to 29.8). Although ca. 62 to 74% of the variation for sex expression was associated with a putative QTL spanning the *F* locus (OP-AJ-2 to *F* and *F* to *de*), other regions (three) of the genome were important for the determination of sex in the F3 families examined depending upon environment. The number of genomic regions affecting main stem length (five) and number of lateral branches (three) coincided with expectations as determined by calculations of minimum number of genes in previous studies. Evaluation of QTL associated with several fruit number determinants of early, first-harvest yield demonstrating additive genetic variance (i.e., sex expression, main stem length, and number of laterals) suggests that marker-assisted selection may have utility for the development of determinate, multiple lateral branching germplasm suited for once-over mechanical harvesting in this population.

Introduction

The commercial cucumber (*Cucumis sativus* L. var. *sativus*; hereafter referred to as *C. s.* var. *sativus*) has a narrow genetic base [6, 16]. Thus, the use of exotic germplasm possessing economically important traits is desirable for population development and introgression of unique genes in this species. For instance, plant habit attributes from a wild-type cucumber, *C. sativus* var. *hardwickii* R. Alef. (hereafter referred to as *C. s.* var. *hardwickii*) have been introduced in commercial cucumber during population development to increase fruit yield in *C. s.* var. *sativus* [37].

Efforts have been made to broaden the genetic base of *C. s.* var. *sativus* germplasm using *C. s.* var. *hardwickii* [24]. Nevertheless, undesirable traits carried along in such wide crosses [20] can delay population improvement efforts, and thus adapted germplasm is

Mention of a trade name, proprietary product, or specific equipment does not constitute a guarantee or warranty by the USDA and does not imply its approval to the exclusion of other products that may be suitable.

often used for insuring short-term advances from selection. Recent development of multiple lateral branching and sequential fruiting genotypes, such as the *C. s.* var. *sativus*-derived inbred line H-19, has provided an opportunity to introgress economically important characteristics in elite cucumber germplasm [11, 32]. H-19 has potential for increasing yield in commercial cucumber because it possesses unique traits that can dramatically alter relationships among yield components (e.g., sex expression, days to flower, and fruit number and weight) in standard indeterminate varieties. When placed in a determinate genetic background such multiple lateral branching cucumber types may have utility in once-over mechanical harvesting operations.

Traditional genetic analyses of diverse cucumber populations have indicated that genetic variance for yield components is directly related to the genetic distance between the parents involved [24, 32, 34, 43]. Likewise, genome mapping of cucumber has indicated that the frequency of polymorphism increases progressively as the genetic distance between the parents increases [14]. Various protein and DNA-based technologies have been evaluated for their potential use in plant improvement [2, 44, 45]. Molecular markers can be used to assist in the introgression of economically important genes during population improvement and to increase the rate of gain from selection [3, 7, 27]. In order to use marker-assisted selection (MAS) effectively in plant improvement programs, relatively saturated linkage maps (i.e., mean distance between markers ca. 5–10 cM) must be constructed and judiciously applied to maximize cost-benefits [41].

Isozymes, restriction fragment length polymorphisms (RFLPs), morphological markers, and more recently random amplified polymorphic DNAs (RAPDs) have been used extensively for genome exploration in cucumber [14, 15]. Each marker type has advantages and disadvantages incumbent for their use [41]. Polymerase chain reaction (PCR) methodologies, such as RAPDs, which amplify DNA fragments in regions containing moderate to high copy sequences can provide for markers in regions of the genome previously inaccessible to analysis by Southern hybridization [25]. Thus, well-defined, repeatable RAPD markers can be used in conjunction with codominant markers to better describe genomes [42]. Where the genetic distance between mapping parents is relatively small (i.e., in narrow crossing), the use of RAPD technologies may provide for more critical elucidation of the genome.

Using the formulae given by Beckman and Soller [2], the genomic length of cucumber can be estimated to be between 800 and 1000 cM. Kennard et al. [14] constructed a 58-point map in a narrow cross and a 70-point map in a wide cross in cucumber to span 766 and 480 cM respectively. While the map constructed from the narrow cross consists of RFLP (31), RAPD (20), isozyme (5), disease resistance (1) and morphological (1) markers, the map developed from the wide cross does not include RAPD markers. RFLPs were used in the narrow cross to investigate QTL associated with fruit quality components [15], and in the wide cross to identify QTL associated with yield components [5]. Since existing cucumber maps are not well saturated and QTL for yield components have not been documented in a narrow cross using molecular markers, there is a need for the characterization of additional molecular loci in cucumber. Therefore, we used line H-19 in a narrow cross in cucumber to construct an 80-point RAPD and morphological markers map and identify QTL associated with yield components for once-over harvest.

Materials and methods

Germplasm and population development

The gynoecious determinate cucumber line G421 possessing normal sized leaves was crossed with the monoecious indeterminate little leaf line H-19. G421 was received from R. L. Lower, University of Wisconsin, and H-19 is a line released by the University of Arkansas, Fayetteville, AK as 'Arkansas Little Leaf'. G421 typically produces 1 to 2 fruit per harvest on 1 or 2 lateral branches originating from the main stem. In contrast, H-19 produces 5 to 20 branches depending upon growing environment, and possesses a sequential fruiting habit (i.e., several fruit enlarge on a lateral branch). The F_1 between G421 $\times$ H-19 was subsequently self-pollinated to produce F_2 progeny and 100 F_3 families.

RAPD analysis

Frozen tissue from 100 F_2 plants was used for DNA extraction. Standard CTAB phenol/chloroform extraction procedures were modified to optimize DNA quality and quantity in a mini-prep format [42]. About 0.1 to 0.3 gm of frozen, pulverized tissue was placed in a 2 ml microcentrifuge tube along with two 4 mm stainless

steel ball bearings, chilled in liquid nitrogen, vortexed for 15 s, rechilled, then vortexed for an additional 15 s. About 750 μl of CTAB buffer (2% CTAB, 1.4 M NaCl, 20 mm EDTA, 0.1 M Tris, 1% 2-mercaptoethanol) was added, vortexed for 30 s, then incubated for 1 h at 60 °C. The mixture was extracted once with phenol/chloroform/isoamyl alcohol (25:24:1) and once with chloroform/isoamyl alcohol (24:1), the ethanol precipitated, and resuspended in 250 μl of 10 mm Tris (pH 8.0) and 0.1 mM EDTA, with 10 units of RNaseI (Promega, Madison, WI). The solution was incubated at 37 °C for 30 min and then stored at −20 °C.

PCR amplifications were performed in 15 μl reaction volumes containing 50 mM KCl, 10 mM Tris (pH 9.0), 0.1% Triton X-100, 3.0 mM $MgCl_2$, 0.2 mM of each of the dNTPs, 0.33 μM 10-mer primer (Operon Technologies, Alameda, CA; University of British Columbia, Vancouver, BC, Canada), 15 ng of genomic DNA, and 1 Unit of *Taq* DNA polymerase (Promega, Madison, WI) [42]. All RAPD reaction products were electrophoresed in 20 cm × 25 cm, 1.6% agarose gels with 0.5 μg/ml of ethidium bromide in 1× TAE buffer. Gels were run for 3 h at 93 V in Gibco/BRL H4 gel apparatus (Bethesda, MD), illuminated by UV light and photographed with an Eagle Eye still video system (Stratagene, La Jolla, CA).

To survey parental genotypes (G421 and H-19) for polymorphisms, banding variation (presence or absence) was recorded for 1520 random 10-mer (Operon (OP-A1–OP-Z20 and OP-AA1–OP-AY20) and British Columbia (BC-200–BC-700)) primers. Putative multiple loci revealed by one primer were designated by a hyphen after the primer array designation (e.g., two loci at OP-B12-1 and OP-B12-2), and their genetic predictability was determined by segregation analysis. Each primer/DNA combination was replicated at least two times in two different PCR runs to confirm segregation in F_2 progeny. Individuals lacking bands were re-analyzed to confirm their absence.

Field evaluation

A total of F_3 families, their parents, and F1 hybrid progeny were evaluated in two U.S. locations: Tifton, Georgia (Ultisol Plintic Paloedults: fine loamy siliceous, thermic) and Hancock, Wisconsin (Typic Udipsament; sandy mixed, mesic). Families were arranged in randomized complete block design with three replications per location. Each replication had fifteen plants, and consisted of single rows with plants spaced 0.7 m apart in rows positioned on 1.5 m centers. This plant density was chosen because it provided a manageable and unbiased appraisal of the variables examined [32].

Characters examined

Data were collected on plant size, days to anthesis, sex expression, leaf size, number of lateral branches originating from the main stem, length of the main stem, fruit number and weight, fruit length and diameter, and fruit length/diameter ratio (L:D) was calculated (Table 1). While the classification of sex expression, and measurements of numbers of lateral branches and main stem length were made on individual plants, data for other traits were taken collectively per replication. Plants were harvested once when ca. 10% of the fruit on plants within a replication were greater than ca. 4.0 cm in diameter (75 and 71 days after planting in Wisconsin and Georgia, respectively).

Determinate (*de*) plant habit was determined by observing plants six weeks after sowing. Plants which possessed short vines with the main stem terminating in a flower cluster were classified as determinate [28].

The number of days to anthesis was taken as the days from planting when at least three plants per replication had reached anthesis.

Sex expression was recorded as a discrete codominant or dominant trait [41], and as a quantitative trait. When interpreted as a discrete character, families were designated as staminate (all male), pistillate (all female) or monoecious (i.e., heterozygous, bearing staminate and pistillate flowers). The proportion of pistillate nodes in the first 10 flower-bearing nodes was used for recording sex expression as a continuous trait. Data were collected on the first ten plants of each replication. A node was considered female if it had at least one pistillate flower.

The number of lateral branches in main stem and main stem length were recorded on the first ten plants on each plot. The first node was taken as that node above the cotyledonary node. Lateral number was taken 2 days before harvest and main stem length was taken at harvest.

Fruit length, diameter, and L:D ratio were determined at harvest by measuring 5 randomly selected fruits from each plot that were greater than 3.5 cm in diameter.

Plants were classified as little, standard or intermediate leaf types by visual examination one week before harvest. Extreme parental leaf types were easily separated by general inspection (little leaf (*ll*) area of the first

Table 1. Means for G421 and H-19, and mean, and heritabilities (h^2) for F_2-derived F3 lines (F3:2) for several horticultural traits in cucumber.

Generation	SEX	LT	DA	NLP	LMS	NFP	FWP	FL	FD	LD
G421	0.96 a	1.0 c	34.3 a	0.96 c	22.7 b	2.1 a	0.95 a	4.8 a	2.18 a	2.26 a
H-19	0.04 d	3.0 a	31.7 b	6.30 a	43.2 a	1.8 a	0.41 b	4.9 a	1.79 c	2.60a
F3:2	0.14 c	1.8 b	32.0 b	3.12 b	35.8 a	2.5 a	1.14 a	5.1 a	1.94 bc	2.61a
lsd	0.04	0.6	0.40	0.25	0.54	0.21	0.18	0.14	0.07	0.07
h^2	0.15*	0.38	0.21	0.52*	0.48*	0.01	0.29	0.11	0.07	0.06

SEX, sex expression; LT, leaz size; DF, number of days to anthesis; NLP, number of laterals/plant; PL, length of mainstem; NFP, number of fruits/plant; FWP, fruit weight/plant; FL, fruit length; FD, druit diameter; LD, fruit length/diameter ratio. * = Narrow sense heritability.

fully expanded leaf = 30 to 40 cm^2, and standard leaf area = 80 to 100 cm^2 [36]). If a family was homogeneous for one leaf type it was considered homozygous (class 1 (little) or 3 (standard)). Where intermediate leaf size phenotypes (leaf area of first fully expanded leaf = ca. 50 to 70 cm^2) were observed in a plot along with parental phenotypes, the family was considered segregating (heterozygous; class 2).

Linkage analysis

Linkage analysis and the order of RAPD loci were estimated using MAPMAKER version 2.0 for Macintosh (based on 21). χ^2 goodness-of-fit tests were performed on putative loci, and those deviating from expected 3:1 segregation ratios at $p < 0.05$ were excluded from linkage analysis. Of the 180 putative polymorphic bands (loci) observed, segregation of 80 bands in 73 primers adequately fit expected ratios in F_2 progeny. These loci were used for map construction and QTL analyses. One hundred bands were eliminated because they were not consistently bright and/or their presence was unpredictable [42].

Putative linkage groups initially were determined by comparisons of pairwise combinations of markers using the 'group' command with a minimum LOD linkage score of 3.0 for statistical acceptance, and minimum recombinant fraction of 0.3. Three-point analyses were performed on each linkage group, taking consecutive markers three at a time. A LOD value of 3.0 was used as the linkage criterion for each triplet combination, and the three-point analysis generated the most likely order for the loci examined. The most likely order for the linkage group was tested with the 'ripple' command. The 'drop marker' command was used to determine whether the presence of individual markers was causing an overestimate of the map size because of experimental errors due to scoring (default setting of 1% a priori probability of error) [23]. When a LOD error ≥ 1.0 was detected, progenies were re-evaluated at least twice and adjustments (corrected or entered as missing data) were made in the data set for those progeny that were found to be inconsistent. The Kosambi mapping function was used to convert recombination fractions to map distances [17].

QTL analysis

Prior to QTL analyses, data of trait means were tested for normality using the PROC UNIVARIATE routine in SAS [29]. Skewness was detected in data for fruit diameter, and therefore logarithmic transformation was performed on mean values to improve normality (log transformed data reported).

Analysis of variance (ANOVA) was performed for all traits to determine the effects of sources of variation (locations, families and interactions). Variance components and heritabilities were previously determined by Serquen et al. [32], and were used for interpreting the outcome of the QTL analyses. One-way analyses of variance ($p < 0.001$) and interval mapping were used to test the association of markers with the QTL for each trait. Both methodologies were used to confirm detection of QTL. The first test was performed using the PROC GLM routine in SAS [29]. Interval mapping tests with a LOD score of 2.0, as the threshold for detecting QTL locations [22], were performed using MAPMAKER/QTL version 1.1. Multi-locus models were constructed using MAPMAKER/QTL and multiple regression via backwards elimination [29]. The latter was used to estimate the magnitude of the phenotypic variance attributable to each QTL (i.e., R^2 presented as %).

Results

A RAPD map: Polymorphisms and segregation of RAPD and morphological markers

Screening of parental stocks with 1520 10-mer primers resulted in the detection of 180 polymorphic bands. Segregation analyses of 73 (4.8%) primers resulted in the identification of 80 loci (i.e., mean of 1.1 marker bands per primer) which were used in map construction. These results are similar to those of Kennard et al. (15), where ca. 5.5% of the RAPD markers examined segregated in a narrow cross (GY-14 × PI 432860).

One primer (OP-AC17) segregated as a codominant marker. Thirteen primers which demonstrated polymorphism between the parents did not segregate in F_2 progeny. This monomorphic condition in F_2 progeny may be due to low amplification levels in some loci or to maternal inheritance. Often repeating the PCR after the initial screening of parental stocks led to the identification of a band in one of the parents where there was formerly no band present.

Five primers (OP-AH14, OP-I20, OP-W7-1, OP-AG17, OP-N8) exhibited deviations from the expected segregation ratios ($p < 0.05$). In four of these primers (all except OP-AH14), distorted segregation was due to an excess of individuals possessing alleles characteristic of the G421 parent. The 'drop marker' command indicated that OP-AH14 caused map expansion (>4.0 cM) when it was added individually to the marker array. Since results from this primer were not repeatable and its banding was difficult to interpret (i.e., light band), it was omitted from the analysis. Three primers (OP-C10, BC 605 and BC 503) could be placed on a linkage group only when the LOD threshold was relaxed below 3.0. These RAPD primers were not placed on the map (i.e., unlinked), but two were significantly associated with QTL for some traits (e.g., fruit weight (BC 605) and fruit number (BC 503)).

Segregations of F_3 progenies at the *F* locus adequately fit predicted ratios for either codominant (1:2:1) or dominant (3:1) classification in each environment tested ($p < 0.05$). Female sex expression in Georgia was greater than in Wisconsin (Table 2). Since significant differences ($p < 0.05$) were observed between environments, mapping was performed separately for each environment. Although the *F* locus mapped to the same linkage group regardless of environment, interval values (mean ± SE) calculated from each location did not span the same position on linkage group B (data not presented). The interval difference between the map position of the *F* locus for the environments evaluated was about 26 cm. This lack of consistency between the placement of the *F* locus resulted in an expansion of the map. Since gynoecy was more frequent in Georgia, only data from this location was used for placement of the *F* locus on the map [30].

A 80-point map consisting of 77 RAPD markers and three morphological markers distributed among nine linkage groups (designated A to I) spanning 599.6 cM was constructed (Figure 1). The average distance between markers was 8.4 ± 9.4 cM. The longest linkage intervals between markers were observed between the *F* locus and OP AJ-2 (44 cM) and the *F* locus and *de* (31.2 cM) in group B. This linkage group was the longest (199 cM) of the nine groups identified and contained 26 (31% of total) markers. Neither the *F* nor the *de* locus were terminal markers on this linkage group (e.g., *de* locus was linked to OP-L18-2, 16.0 cM). The *ll* locus mapped to linkage group D and was flanked by BC-551 (3.4 cM) and BC-592 (12.2 cM).

QTL Analysis

Analyses of variance performed on F_3 family means detected significant differences between locations (L), and among families (F), and interactions between families and locations for most of the traits examined (Table 2). For leaf size, days to anthesis, number of lateral branches, length of main stem, and number of fruits per plant the L × F interaction was not significant ($p < 0.05$). Nevertheless, since most of the traits examined exhibited L × F interaction, the QTL analysis was performed on a location basis.

Results obtained from one-way ANOVA ($p < 0.001$) and MAPMAKER/QTL (LOD score 2.0) analyses were essentially equivalent (i. e., analyses defined similar marker-trait associations). Increasing p values (i.e., decreasing stringency) in one-way ANOVAs (e.g. $p < 0.05$, $p < 0.01$, for individual markers) resulted in additional genomic regions showing significant associations with most traits. These QTL may either represent minor genetic factors affecting a trait's expression or spurious associations. The threshold significance level for marker-trait associations used was conservative ($p < 0.001$) in order to reduce the probability of false-positives [22]. Using MAP MAKER/QTL, 16 genomic regions were found to be associated with the traits examined in Georgia, and 22 regions associations were detected in Wisconsin (Table 3). Sex expression, number of lateral branches and main stem length exhib-

Table 2. Analysis of variance of G421 x H19-derived F3 lines (F3:2) for several horticultural traits in cucumber averaged over three replications and two locations.

Source	df	SEX	LT	DA	NLP	LMS	NFP	FWP	FL	FD	LD
Location (Loc)	1	0.431**	0.6215	2574.6**	0.0069	14.53**	6.752**	3.914**	4.574**	3.320**	0.056**
Rep/Loc	4	0.012**	0.0718	15.7**	0.6991**	0.59	1.595**	0.922**	0.153**	0.025**	0.083**
Families (Fam)	103	0.008**	0.2723	9.5**	0.4046**	2.07**	0.125**	0.163**	0.083**	0.024**	0.014**
Loc × Fam	96	0.0018*	0.0628	5.8	0.0770**	0.39*	0.088	0.065**	0.058**	0.019**	0.011**
Error		0.0013	0.0667	4.047	0.0801	0.30	0.078	0.040	0.026	0.007	0.007

SEX, sex expression; LT, leaf size; DA, number of days to anthesis; NLP, number of laterals/plant; LMS, length of mainstem; NFP, number of fruits/plant; FWP, fruit weight/plant; FL, fruit length; FD, druit diameter; LD, fruit length/diameter ratio.

ited the largest number of QTL of the traits examined. This might have been expected given the differences observed between the parents for these traits (Table 1).

Sex expression. Five QTL for sex expression were detected in Georgia using MAPMAKER/QTL, four were identified in Wisconsin (Table 3). The unique QTL for sex expression (spanning OP-H5 to OP-R13) detected in Georgia was located on linkage group C (Figure 1). The largest QTL for sex expression was associated with the *F* locus in linkage group B (OP-AJ-2 to *F* LOD = 28.3 in Georgia, and *F* to *de* , LOD = 29.8 in Wisconsin).

Main stem length. Three QTL for main stem length were detected in each location (Table 3; Figure 1). The major contribution to variation in main stem length was detected in the interval *F-de*, which explained about 45% and 39% of the total variation for this trait in Georgia and Wisconsin, respectively. Additional chromosomal regions which defined variation for this trait were found in linkage groups C (OP-R13 to OP-H5; $R^2 = 6.5\%$) and E (OP-U15-2 to OP-I20; $R^2 = 8.0\%$) in Georgia, and groups D (*ll* to BC-592; $R^2 = 12\%$) and G (OP-AB14 to BC-469; $R^2 = 9.5\%$) in Wisconsin.

Number of lateral branches per plant. A common QTL (spanning *de* to OP-L18-2) was identified for both locations on linkage group B (Table 3; Figure 1). This QTL explained about 40% and 37% of total phenotypic variance in Georgia and Wisconsin, respectively. Additional QTL were detected in linkage group D that also explained the variation observed in Georgia (BC-403 to OP-W7-2, R^2 = ca. 14%) and Wisconsin (OP-AJ6 to BC-523, $R^2 = 11\%$).

Days to anthesis. Two location-dependent QTL were detected for days to anthesis on linkage group B (Table 3; Figure 1). While one QTL (spanning *de* to OP-L18-2, R^2 = ca. 13%) was identified from phenotypic variation in Georgia, two QTL (spanning OP-C7 to BC526, R^2 = ca. 8.0 and OP-L18-2 to OP-J5-1, R^2 = ca. 13.0) were detected in Wisconsin.

Fruit number. One region on linkage group B (spanning *de* to OP-L18-2, R^2 = ca. 9.0%) was found to be associated with variation in fruit number in Georgia (Table 3, Figure 1). In Wisconsin, a region on linkage group B (spanning *F* to *de*) also explained about 9% of the variation for this trait. In addition, a region on linkage group D (spanning *ll* to BC-592) and a region on group H (spanning OP-M8-1 to OP-N6-2) explained ca. 20% and 6.0% of the variation, respectively.

Fruit weight. One region containing the *ll* locus was detected in linkage group D (BC-551 to *ll*) which explained a major portion of the variation for fruit weight in Georgia (R^2 =ca. 40%) (Table 3, Figure 1). About 38% of the variation for this trait was explained by three QTL in Wisconsin (group B, L18-2 to OP-J5-1; group D, *ll* to BC-592; group H, L18-2 to OP-J5-1).

Fruit length.

One region on linkage group D (spanning *ll* to BC-592) was detected that explained a major portion of the variation for fruit length in Georgia (R^2 = ca. 21%) and Wisconsin (R^2 = ca. 31%) (Table 3, Figure 1).

Fruit diameter.

Two regions, one on linkage group B (spanning *de* to OP-L18-2) and another on linkage group D (spanning BC-551 to *ll*) collectively explained ca. 38% of

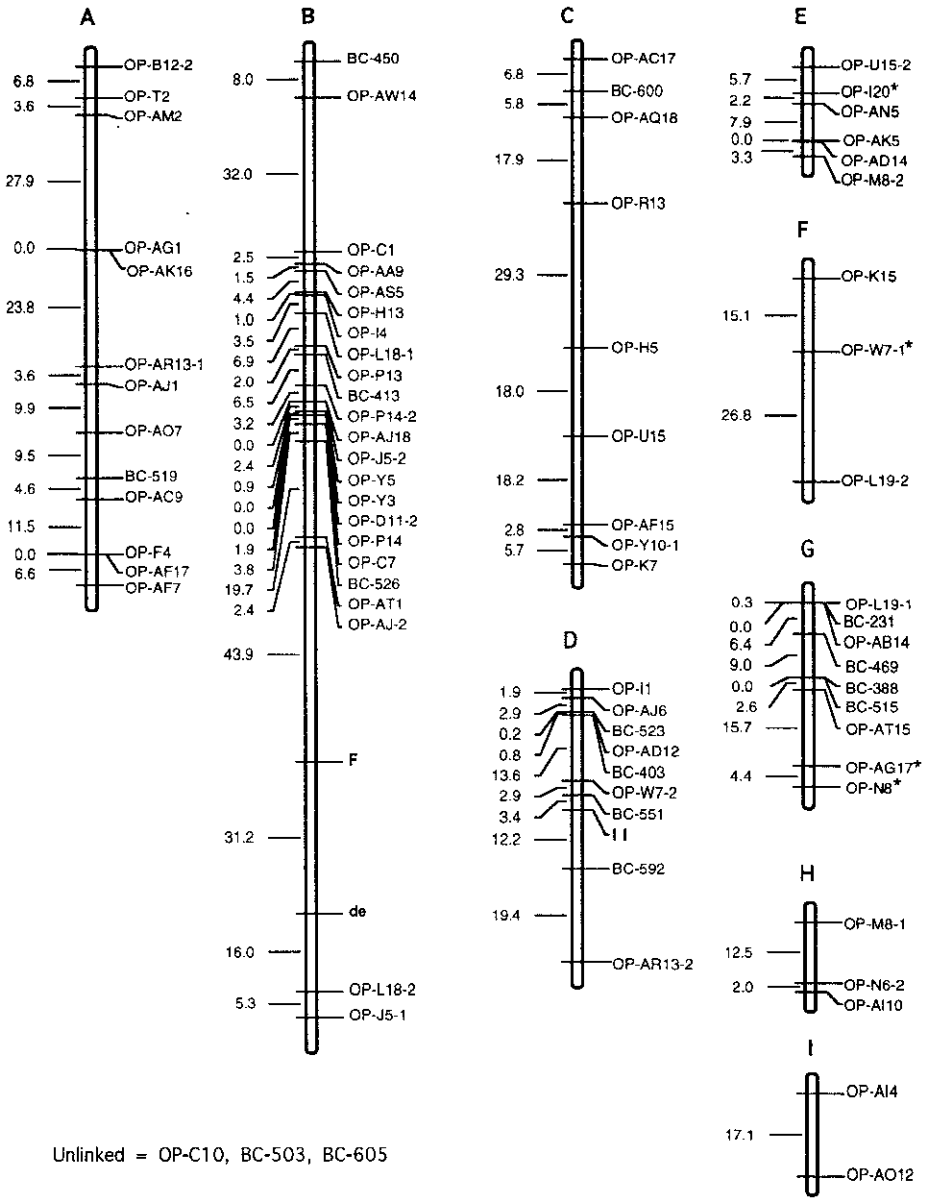

Figure 1. RAPD-based linkage map of cucumber (*Cucumis sativus* L. var. *sativus*). Linkage groups are designated by upper-case letters, morphological markers are in bold, and RAPD markers are named as described in Materials and methods. Aberrantly segregating loci are designated as * ($P < 0.05$).

the variation for fruit diameter in Georgia (Table 3, Figure 1). Likewise, two regions identified on linkage group B (spanning OP-C7 to BC-526 and *F* to *de*) explained ca. 19% of the variation for this trait in Wisconsin.

Length/diameter ratio. Two regions, one on linkage group B (spanning *F* to *de*) and the other on linkage group D (BC-551 to *ll*) were found to explain 28% of the variation for length:diameter ratio only in Wisconsin (Table 3, Figure 1).

Table 3. Significant marker QTL associations detected by with interval mapping in a narrow cross in cucumber.

Trait	Interval	Linkage group	Georgia		Wisconsin	
			LOD	R^2 (%)	LOD	R^2 (%)
Sex expression	OP-AW14 - OP-C1	B	4.9	10.4	5.08	8.1
	OP-AJ-2-*F*	B	28.3	67.4		
	F -de	B			29.8	74.2
	OP-H5-OP-R13	C	2.7	7.8		
	OP-W7-1-OP-L19-2	F	4.1	6.0	3.4	5.2
	OP-AI4-OP-AO12	I	3.0		2.1	
Mainstem length	*F-de*	B	12.4	44.5	13.4	39.0
	OP-R13-OP-H5	C	2.5	6.5		
	ll-BC-592	D			2.4	12.1
	U15-2-OP-I20	E	2.4	8.0		
	OP-AB14-BC-469	G			2.4	9.5
No. of laterals	*F*	B	5.5	13.1		
	de-OP-L18-2	B	10.4	39.6	10.1	37.0
	BC-403-OP-W7-2	D	4.6	13.6		
	OP-AJ6-BC-523	D			3.3	11.0
Days to anthesis	*de*-OP-L18-2	B	3.9	12.8		
	OP-C7-BC-526	B			3.1	8.1
	OP-L18-2 - OP-J5-1	B			3.4	12.9
Fruit number	*de*-OP-L18-2	B	2.1	9.3		
	F-de	B			2.7	9.2
	ll-BC-592	D			4.2	19.8
	OP-M8-1-OP-N6-2	H			2.4	6.0
Fruit weight	BC-551-*ll*	D	11.9	39.7		
	OP-L18-2-OP-J5-1	B			2.8	9.2
	ll-BC592	D			4.2	20.6
	OP-M8-1-OP-N6-2	H			2.5	7.8
Fruit length	*ll* -BC-592	D	6.1	21.4	6.1	31.0
Fruit diameter	*de*-OP-L18-2	B	3.7	15.7		
	OP-C7-BC-526	B			2.6	9.2
	F-de	B			2.8	9.6
	BC-551-*ll*	D	4.7	21.9		
Ratio L/D	*F-de*	B			3.3	13.7
	BC-551-*ll*	D			3.1	14.4

Discussion

The acreage for once-over mechanical harvesting of processing cucumber is increasing in U.S. northern production areas [40]. Thus, there is a need to develop high yielding varieties for once-over mechanical harvesting. Determinate multiple branching genotypes provide potential for increasing yield in cucumber [39]. The discrete nature of *de* and the additivity of genes for lateral branching and main stem length suggest that the selection for determinate multiple lateral phenotypes differing in size is possible using the narrow cross in this study [32]. However, vegetative propagation of determinate genotypes is difficult and gentoype mis-

classification of determinate multilateral phenotypes due to a lack of apical dominance makes selection inherently difficult [39]. Thus, genetic markers that allow for early selection of determinate genotypes in a multiple lateral background would be a useful addition to traditional phenotypic selection strategies.

The linkage map reported herein is the second to utilize RAPDs in cucumber [14] and consists of nine linkage groups. Although the map was constructed principally with RAPD markers, it defines the relative map location of *de*, *F*, and *ll*, and QTL associated with yield components in a narrow cross. Its total length is shorter by comparison than a map derived from a narrow cross reported by Kennard et al. [14] (ca. 600 vs. ca. 766 cM), but longer than a map developed from a wide cross (ca. 480 cM) in cucumber. Given current estimates of total length [14], this map spans 85% of the genome with an average distance between markers of 8.4 ± 9.4 cM. Some regions of this map contain gaps and others are well saturated. For instance, relatively large gaps (spanning OP-AJ-2 to *F* and *F* to *de*) and a cluster (spanning OP-C1 to OP-AJ-2) of markers was observed on linkage group B (Figure 1). Such gaps have been found in genetic maps of other species (e.g., [26]) and does not necessarly preclude QTL analysis and MAS [18]. Although the addition of markers may fill existing gaps, such additions will not necessarily enhance QTL detection or map utility in MAS [10].

The maps of Kennard et al. [14] defined 10 linkage groups in both narrow and wide crosses in cucumber. This is greater than the seven predicted for this species and indicates the existence of gaps in genomic information in these maps. Using the data from segregating progeny in this study, a complement of seven chromosomes was achieved when the LOD score was relaxed to 2.96 and an *r* of 0.3 (linkage groups not presented). When these thresholds were employed, linkage groups B and E, and groups F and H merged to form two larger groups. Relaxing the stringency, however, increases the chance of identifying spurious linkages which results in type I errors.

Fruit number determinants of early, first-harvest yield in cucumber involves sex expression type and source sink relationships (e.g., main stem length and multiple branching habit) [32]. Genetic estimates of variance and covariance for yield components in this mapping population (G421 $\times$ H-19) have been examined [32]. Additive and dominant genetic variances were important for the expression of sex. Differences among F_3 families in the expression of gynoecy could be explained by a lack of complete dominance at the *F* locus [33], or by the proposed existence of multiple *F* alleles [19]. Main stem length and multiple lateral branching exhibited mostly additive genetic variance. The minimum number of genetic factors controlling number of lateral branches and main stem length were estimated to be 4, and 8, respectively [32].

The data presented provide support for the control of gynoecy at the *F* locus [14, 33], and evidence for the role of other chromosome regions in determining sex in cucumber (Table 3). The magnitude of QTL associated with *F* locus (linkage group B; Figure 1) detected in this study, shows the relative importance of this locus in determining gynoecious sex expression in cucumber (Table 3). The *F* locus was also mapped to a unique position in other maps constructed from narrow and wide crosses in cucumber [14]. However, in all previously constructed cucumber linkage maps [14] and the one created in this study, genetic markers were positioned relatively far from the *F* locus. The closest marker (RFLP) was positioned 9 cM from the*F* locus in a wide cross (*C. s.* var. *sativus* GY14 $\times$ *C. s.* var. *hardwickii* PI 183567) [14].

Studies by Trebitsh et al. [46] indicate that a genomic region encoding ACC synthase is located in linkage group B coincident with the *F* locus (recombination frequency of 0%). This region is either the *F* locus itself or a closely linked female modifying factor. In a wide cross of cucumber, 90% of the variation for sex expression was explained by a QTL which included the *F* locus [14]. In our study, ca. 67% and ca. 74% of the variation for sex expression was associated with one QTL and included the *F* locus. In both studies, data indicate that other regions of the genome interact to modify sex expression. It is plausible that, in our population, the genomic regions identified on linkage groups B (OP-AW14 to OP-C1), C (OP-H5 to OP-L19-2), F (OP-W7-1 to OP-L19-2), and I (OP-AI4 to OP-AO12) (Table 3) may operate to modify the action of the *F* locus in cucumber.

It is well known that trait expression can be affected by environment [1, 8]. The fact that QTL of comparatively smaller magnitude in this study were not consistently expressed across locations (e.g., sex expression QTL detected in Georgia spanning OP-H5 to OP-R13) may be related to genotype by environment interaction (Tables 2 and 3). The difference in sex expression between Georgia and Wisconsin was predictable since expression of the *F* locus is known to be affected by environment [4, 9] and sex-modifying genes [19, 48]. The ambient temperature, light intensity and relative humidity was higher in Georgia when compared to

Wisconsin (data not presented). These and other factors undoubtedly played a role in the determination of the sex expression differences observed between growing locations (Table 2). In contrast, although determinate plant habit (*de*) and leaf size (*ll*) can also be affected by environment [36, 38, 39], their expression across environments in this study was relatively uniform, and therefore these characters mapped uniformly to unique map positions irrespective of growing environment (B and D, respectively, Figure 1).

The number of genomic regions affecting main stem length (five) and number of lateral branches (three) were relatively low and close to the expectations as determined by computational estimates of numbers of genetic factors [32]. Genetic effects (i.e., dominance, additivity, epistasis) of QTL are difficult to estimate, and are often underestimated [31], because genotypes can not be completely classified during RAPD analysis. The genomic regions affecting main stem length were fewer than the estimated genetic factors because some putative genomic regions were eliminated due to the mapping thresholds used. Moreover, since elite parents were used for mapping, some horticulturally important traits (e.g., yield components) may have been fixed in the inbred lines and thus the QTL detected may represent small regions affecting traits of interest.

Using RFLP mapping in a narrow cross of cucumber, Kennard and Havey [15] found five QTL to be associated with fruit length and three QTL to be associated with fruit diameter. Isozyme and RFLP analysis in a wide cross of cucumber detected six QTL associated with fruit length and three associated with fruit diameter [5]. Our analysis using RAPDs in a narrow cross identified three QTL for fruit diameter and one for fruit length. Although the QTL cannot be compared directly, these data suggest that the number of QTL controlling fruit dimensions in cucumber is relatively low. This hypothesis is supported by the fact that fruit length and diameter in cucumber can be manipulated relatively easily by various phenotypic selection strategies.

The principal objective for mapping and QTL analysis in our laboratory is to evaluate MAS for increasing gain from selection in cucumber. Adjustments in plant architecture in cucumber can provide opportunities to increase yield. Plant measurements in this study were taken at relatively wide spacing (ca. 19 000 plants/ha). Genotypes used in once-over machine harvest operations are typically placed at plant densities between 150 000 to 200 000 plants/ha. The performance of genotypes identified as having utility for once-over machine harvest at low plant densities may change when they are placed at high densities. Nevertheless, rankings of relative yield performance of indeterminate, multiple branching genotypes derived from *C. s.* var. *hardwickii* were similar at different planting densities [5], and determinate, multiple-branching genotypes have been identified which retain yield performance at high densities [38]. Thus, MAS may have potential for identifying unique, high yielding genotypes in this and genetically similar populations.

Molecular approaches for crop improvement are demonstrating potential for enhancing single gene disease resistance in several species [13, 47]. An important challenge is to clarify the potential usefulness of MAS for complex traits whose expression(s) is often the result of complex relationships among genetic factors, environmental factors, and their interactions. Our data suggest that MAS for main stem length and multiple lateral branching may be effective. Because of the LOD thresholds used ($>p = 0.001$) for QTL identification, it is likely that the QTL identified in this study would have value for MAS in other populations derived from H-19 if PCR conditions are stringently controlled [42]. However, the cost of MAS and the fact that a close linkage between a genetic marker and the *de* locus has not been found (i.e., 16 cM between OP-L-16-2 and *de*; Figure 1) are important considerations for using MAS in this population to identify determinate multiple lateral branching types.

It would be of benefit to increase map saturation in cucumber. The addition of markers may allow for the detection of closer linkages between markers and *de*, and may increase the frequency of markers associated with QTL regions. The merging of existing cucumber maps may afford an avenue for developing a more saturated, unified map in cucumber. Such map merging could be accomplished by the examination of common markers among appropriate mapping populations and the use of sophisticated computer software [12, 35]. Alternatively, other marker techniques could be used to to identify additional polymorphisms in selected mapping populations in an attempt to insert markers into existing gaps in these maps [41].

References

1. Allard RW: Principles of Plant Breeding. John Wiley, New York/London (1960).
2. Beckmann JS, Soller M: Restriction fragment length polymorphisms in genetic improvement: methodologies, mapping, and costs. Theor Appl Genet 67: 35–43 (1983).
3. Bonierbale MW, Plaisted RL, Pineda O, Tanksley SD: QTL analysis of trichome-mediated insect resistance in potato. Theor Appl Genet 87: 973–987 (1994).
4. Cantliffe DJ: Alteration of sex expression in cucumber due to changes in temperature, light intensity, and photoperiod. J Am Soc Hort Sci 106: 133–136 (1981).
5. Dijkhuizen A: Application of restriction fragment length polymorphism for the assessment of genetic variability and study of quantitatively inherited traits in cucumber (*Cucumis sativus* L.). Ph.D. thesis, University of Wisconsin-Madison (1994).
6. Dijkuizen A, Kennnard WC, Havey MJ, Staub JE: RFLP variability and genetic relationships in cultivated cucumber. Euphytica 90: 79–87 (1996).
7. Edwards MD, Stuber CW, Wendel JF: Molecular-marker-facilitated investigations of quantitative-trait loci in maize. I. Numbers, genomic distribution, and types of gene action. Genetics 116: 113–125 (1987).
8. Falconer DS: Introduction to quantitative genetics. Longman Scientific and Technical, UK (1989).
9. Galun E: Study of the inheritance of sex expression in the cucumber. The interaction of major gene with modifying genetic and non-genetic factors. Genetica 13: 1–8 (1971).
10. Gimelfarb A, Lande R: Marker-assisted selection and marker-QTL associations in hybrid populations. Theor Appl Genet 91: 522–528 (1995).
11. Goode MJ, Bowers JL, Morelock TE: Arkansas little leaf cucumber. American Society of Horticultural Science, Annual Meeting/programs and abstracts (1989).
12. Hauge BM, Hanley SM, Cartinhour S, Cherry JM, Goodman HM, Koornneef M, Stam P, Chang C, Kempin S, Medrano L, Meyerowitz EM: An integrated genetic/RFLP map of the *Arabidopsis thaliana* genome. Plant J 3: 745–754 (1993).
13. Kelly J: Use of random amplified polymorphic DNA markers in breeding for major resistance to plant pathogens. HortScience 30: 461–465 (1995).
14. Kennard WC, Poetter K, Dijkhuizen A, Meglic V, Staub JE, Havey MJ: Linkages among RFLP, RAPD, isozyme, disease resistance, and morphological markers in narrow and wide crosses of cucumber. Theor Appl Genet 89: 42–48 (1994).
15. Kennard WC, Havey MJ: Quantitative trait analysis of fruit quality in cucumber; QTL detection, confirmation, and comparison with mating-design variation. Theor Appl Genet 91: 53–61 (1995).
16. Knerr LD, Staub JE, Holder DJ, May BP: Genetic diversity in *Cucumis sativus* L. assessed by variation at 18 allozyme coding loci. Theor Appl Genet 78: 119–128 (1989).
17. Kosambi DD: The estimation of map distance from recombination values. Ann Eugen 19: 172–175 (1994).
18. Knapp SJ: Mapping quantitative trait loci. In: Phillips RL and Vasil IK (eds) DNA-Based Markers in Plants, pp. 5896. Kluwer Academic Publishers, Dordrecht, Netherlands (1994).
19. Kubicki B: Investigations on sex determination in cucumber (*Cucumis sativus* L.). Genet Pol 10: 3–143 (1969).
20. Kupper R, Staub JE: Combining ability between lines of *Cucumis sativus* L. and *Cucumis sativus* var. *hardwickii* (R.) Alef. Euphytica 38: 197–210 (1988).
21. Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newberg L: MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1: 174–181 (1987).
22. Lander ES, Botstein D: Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121: 185–199 (1989).
23. Lincoln SE, Lander ES: Systematic detection of errors in genetic linkage data. Genomics 1: 174–181 (1992).
24. Lower RL, Nienhuis J, Miller CH: Gene action and heterosis for yield and vegetative characteristics in a cross between a gynoecious pickling cucumber inbred and a *Cucumis sativus* var. *hardwickii* line. J Am Soc Hort Sci 107: 75–82 (1982).
25. Michelmore RW, Paran I, Kesseli RV: Identification of markers linked to disease resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions using segregating populations. Proc Natl Acad Sci USA 88: 9828–9832 (1991).
26. Nodari RO, Tsai SM, Guzman P, Gilbertson RL, Gepts P: Toward an integrated linkage map of common bean. III. Mapping genetic factors controlling host-bacteria interactions. Genetics 134: 341–350 (1993).
27. Paterson AH, Deverna JW, Lanini B, Tanksley SD: Fine mapping of quantitative trait loci using a selected overlapping recombinant chromosome in an interspecies cross of tomato. Genetics 124: 735–724 (1990).
28. Pierce LK, Wehner TC: Review of genes and linkage groups in cucumber. HortScience 25: 605–615 (1990).
29. SAS: SAS/STAT Users guide, release version 6.03. SAS Institute, Cary, NC (1988).
30. Serquen FC, Staub JE: Environment can effect the placement of discrete traits on genetic maps: the case of sex expression in cucumber. Cucurbit Genet Coop Rpt 9: 10–12 (1996).
31. Serquen FC, Staub JE: Underestimation of detection of QTL when using dominant genetic markers: the case of RAPDs in cucumber. Cucurbit Genet Coop Rpt 19: 13–16 (1996).
32. Serquen FC, Bacher J, Staub JE : Genetic analysis of yield components in cucumber (*Cucumis sativus* L.) at low plant density. J Am Soc Hort Sci, in press (1997).
33. Shifriss O: Sex control in cucumbers. J Hered 52: 5–12 (1961).
34. Smith OS, Lower RL, Moll RH: Estimates of heritabilities and variance components in pickling cucumber. J Am Soc Hort Sci 103: 222–225 (1978).
35. Stam P: Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J 3: 739–744 (1993).
36. Staub JE, Knerr LD, Hopen HJ: Effects of plant density and herbicides on cucumber productivity. J Am Soc Hort Sci 117: 48–53 (1992).
37. Staub JE, Peterson CE, Crubaugh LK, Palmer MJ: Cucumber population WI 6383 and derived inbreds WI 5098 and WI 5551. HortScience 27: 1340–1341 (1993).
38. Staub JE, Crubaugh L: Selection for multiple lateral determinate cucumber genotypes. Cucurbit Genet Coop Rpt 18: 5–6 (1995).
39. Staub JE, Bacher J, Crubaugh L: Problems associated with the selection of determinate cucumber (*Cucumis sativus* L.) plant types in a multiple lateral background. Cucurbit Genet Coop Rpt 18: 7–9 (1995).
40. Staub JE, Bacher J: Cucumber as a processed vegetable. In: Processing vegetables: science and technology. Technomic Publishing Company, Lancaster, PA (in press).

41. Staub JE, Serquen FC, Gupta M: Genetic markers, map construction, and their application in plant breeding. HortScience 31: 729–741 (1996).
42. Staub JE, Bacher J, Poetter K: Sources of potential errors in the application of random amplified polymorphic DNAs in cucumber. HortScience 31: 262–266 (1996).
43. Strefler MS, Wehner TC: Estimates of heritabilities and genetic variances of three yield and five quality traits in three fresh-market cucumber populations. J Am Soc Hort Sci 111: 599–605 (1986).
44. Tanksley SD, Medina-Filho H, Rick CM: Use of naturally-occurring enzyme variation to detect and map genes controlling quantitative traits in an interspecific backcross of tomato. Heredity 49: 11–25 (1982).
45. Tansksley S: Molecular markers in plant breeding. Plant Mol Biol Rep 1: 3–8 (1983).
46. Trebitsh T, Staub JE, O'Neil SD: Identification of an 1-aminocyclopropane-1-carboxylate synthase gene linked to the female gene (*F*) that determines female sex expression in cucumber (*Cucumis sativus* L.) Plant Physiol, in press (1997).
47. Young ND, Zamir D, Ganal MW, Tanksley SD: Use of isogenic lines and simultaneous probing to identify DNA markers tightly linked to the Tm-2a gene in tomato. Genetics 120: 579–585 (1988).
48. Zhang Q, Gabert AC, Baggett JR: Parents and mating systems affect the transfer of gynoecious flowering to Chinese monoecious cucumbers. J Am Soc Hort Sci 117: 515–517 (1992).